Office de la propriété intellectuelle du Canada

Un organisme d'Industrie Canada

Canadian Intellectual Property Office

An Agency of Industry Canada

PCT / CA 00/01027

0 T 2000(02.10.00)

CA 00/01027 10/070503

Bureau canadien des brevets

Certification

Canadian Patent Office

Certification

La présente atteste que les documents ci-joints, dont la liste figure ci-dessous, sont des copies authentiques des documents déposés au Bureau des brevets.

This is to certify that the documents attached hereto and identified below are true copies of the documents on file in the Patent Office.

Specification and Drawings, as originally filed, with Application for Patent Serial No: **2,282,179,** on September 7, 1999, by **NOVOPHARM BIOTECH INC.,** assignee of William Herman, Howard Kaplan, Joycelyn Entwistle, Jamshid Tanha, Saran Narang and Michael Dan, for "Enhanced Phage Display Library".

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Agent certificateur/Certifying Officer

October 2, 2000

Date

Canada

(CIPO 68)

OPIC CIPO

ENHANCED PHAGE DISPLAY LIBRARY

Field of the Invention

The present invention relates to the field of molecular immunology, and more particularly, to a phage display library in which the recombinant phage population displays a plurality of potential binding-fragments, having preferred characteristics of solubility and/or intermolecular interaction

Background of the Invention

Recent developments in antibody engineering and recombinant DNA technology have made it possible to generate forms of recombinant antibody fragments that, in many ways, are functional substitutes of larger intact immunoglobulin molecules. Moreover, it is believed that smaller fragments of antibodies are advantageous for pharmaceutical applications, particularly in cancer targeting and imaging to assist in penetration into large solid tumours. In the recent past, single domain heavy chain antibody fragments have been the subject of several reports in the patent and scientific literature. Furthermore, the literature reports efforts to generate phage display libraries of such fragments for biopanning against a target ligand.

U.S. Patent No. 5,702,892 ('892) discloses a phage display library constructed in an M13 derived expression vector, in which recombinant phage of the library contain a polynucleotide encoding a fusion protein which comprises a phage coat protein and an immunoglobulin heavy chain binding-fragment. The heavy-chain binding-fragment spans from a position upstream of CDR1 to a position downstream of CDR3. '892 describes that the DNA sequence encoding the CDR3 region and/or the CDR1 region may be randomly varied so that the population of phage expresses a series of potential heavy chain binding domains for panning against the target ligand.

It is known that camelid heavy chain antibodies occur naturally as heavy chain dimers. Heavy chain antibodies specific for a target antigen may be generated by immunizing a member of the camelid species with the target antigen (see Lauwereys et al. (1998) The EMBO J. 17, 3512-3520).

U.S. Patent No. 5,759,808 discloses a phage display library comprising a population of phage based on random variation of a cDNA sequence obtained from lymphocytes of camelids previously immunized with target antigens. However, an approach requiring immunization has limitations with respect to many types of antigens due to ethical considerations relating to experimentation with these animals.

Hamers-Casterman et al. (1993) Nature 362, 446-448 report that camelid heavy chain antibodies are naturally more hydrophilic at amino acid residues at locations 44, 45 and 47 (Kabat

numbering system), in FR2, which corresponds to the surface where they normally contact the V_L domain. Another salient feature of a camelid V_{HH} is that it generally has a comparatively longer CDR3 with a high incidence of cysteines and thus may form, via paired cysteines in CDR1 and CDR3, exposed loops, which are more amenable to binding into cavities such as the active site of enzymes and antibodies (Desmyter et al. (1996) Nat. Struct. Biol. Vol. 3, No. 9, p. 803). However, it has been questioned whether single domain antibodies with desired affinities can be generated with such configurations in the absence of prior immunization, i.e. with a naïve library (Lauwereys et al. (1998) supra).

There remains a need in the art for a naïve phage display library based on an antibody fragment which is naturally highly soluble; a phage display library which is preferably engineered for preferred intermolecular interactions, built on a natural human V_H scaffold, which is non-immunogenic for human therapeutic use, and naturally has a long CDR3 that is amenable to forming exposed loops for various applications, including binding into cavities.

Summary of the Invention

In one broad aspect, the invention is directed to a population of genetic packages having a genetically determined outer surface protein including genetic packages which collectively display a plurality of potential binding-fragments in association with said outer surface protein, each package including a nucleic acid construct coding for a fusion protein which is at least a

portion of said outer surface protein and a variant of at least one soluble parental binding-fragment, preferably a V_H binding-fragment spanning from upstream of CDR1 to a position downstream of CDR3, wherein at least part of said construct, preferably including at least part of the CDR3, is only partly randomized in that it is biased in favour of encoding the amino acid constitution of said parental binding-fragment such that said plurality of different potential heavy chain binding domains are, on the whole, adapted to be, or better capable of, being expressed as soluble proteins.

In a preferred embodiment, the soluble parental binding-fragment is a V_H binding-fragment comprising a nucleic acid coding sequence spanning from upstream of an immunoglobulin V_H CDR1 to a position downstream of the CDR3, more preferably including at least part of FR1 to and at least part of FR4. By biasing the amino acid constitution, preferably on an individual amino acid by amino acid basis, in favour of the wild-type amino acid constitution, portions of the parental binding-fragment that are randomized in favour of generating variability in the potential binding-fragments can be engineered to maintain favourable solubility characteristics of the parental binding domain. Preferably, a portion of said construct encoding at least part of the CDR3 is biased in favour of the wild-type amino acid constitution. In a further preferred embodiment, the parental V_H binding-fragment naturally has a long CDR3 that is amenable to forming exposed loops for binding into cavities. In a further preferred embodiment, the parental V_H binding-fragment is built on a human framework or is adapted from or adaptable to a human framework.

Accordingly, in a most preferred embodiment, the invention is directed to a population of genetic packages having a genetically determined outer surface protein including genetic packages which collectively display a plurality of potential binding-fragments in association with said outer surface protein, each package including a nucleic acid construct coding for a fusion protein which is at least a portion of said outer surface protein and a variant of at least one soluble parental binding-fragment derived from or having the amino acid sequence identified in SEQ. ID. 1 (or a sequence at least 80% homologous thereto), wherein at least part of said construct, preferably including at least part of the CDR3 identified in SEQ. ID. 1, encodes or is biased in favour of encoding, the amino acid constitution of said parental binding-fragment such that said plurality of different potential heavy chain binding domains are on the whole adapted to be or better capable of being expressed as soluble proteins.

In another aspect, the invention is directed to a parental binding domain that is at least partially biased to have amino acids that are preferred for intermolecular interactions selected from the group consisting of tyrosine, histidine, glutamine, asparagine, lysine, aspartic acid and glutamic acid. In a related aspect the invention is directed to a method of biasing amino acid residues to the previously specified amino acid residues by at least partially biasing the second nucleotide position in a desired coding triplet to adenosine.

Accordingly, in another embodiment, the invention is directed to a population of genetic packages having a genetically determined outer surface protein including genetic packages

which collectively display a plurality of potential binding-fragments in association with said outer surface protein, each package including a nucleic acid construct coding for a fusion protein, which is at least a portion of said outer surface protein and a variant of at least one parental binding-fragment, wherein at least part of said construct, preferably including at least part of the CDR3 is at least partially biased in favour of encoding amino acids selected from the group consisting of tyrosine, histidine, glutamine, asparagine, lysine, aspartic acid and glutamic acid such that said plurality of different potential binding fragments are on the whole better adapted for intermolecular interactions.

It is to be understood that potentially suitable genetic packages include cells, spores and viruses (see US Patent No. 5,571,698), namely replicable genetic packages.

The invention is directed to a heterogeneous population of replicable genetic packages which collectively display a plurality of different V_H binding-fragments, each of said genetic packages being genetically alterable and having an outer surface including a genetically determined outer surface protein. Each replicable genetic package includes a nucleic acid construct coding for a fusion protein which comprises at least a portion of said outer surface protein and a potential V_H binding-fragment preferably spanning from upstream of an immunoglobulin V_H CDR1 to a position downstream of CDR3, said nucleic acid construct comprising a coding sequence, which is biased to conserve part of the amino acid constitution of said parental V_H binding-fragment at least part of which is a randomly or partially randomly generated variant of a segment of DNA encoding a CDR3 of a soluble non-camelid type parental V_H binding-fragment. The result is that

said heterogeneous population of replicable genetic packages collectively encode fusion proteins having a plurality of different potential V_H binding-fragments related to said parental V_H binding-fragment. The fusion proteins are expressed in the absence of an immunoglobulin light chain protein or portions thereof on an outer surface of said genetic packages. The potential V_H binding-fragments are, on the whole, thereby adapted to be or better capable of being expressed as soluble proteins and are preferably characterized by a CDR3 having 16 to 33 amino acids.

Preferably, the replicable genetic package is a recombinant phage and said heterogeneous population of replicable genetic packages collectively constitute a phage display library.
In another embodiment, the parental binding-fragment is a natural occurring antibody having a natural human V_L interface. In another embodiment, the V_L interface is engineered to avoid hydrophobic amino acids. In another embodiment, the V_L interface is engineered for amino acids, which form weak interactions. In its broader aspects, the invention excludes genetic packages encoding potential binding-fragments that have preferred characteristics in virtue of camelid type or engineered V_L interfaces.

Accordingly, the invention is preferably directed to a phage-display library, comprising a plurality of recombinant phage, each of said recombinant phage having incorporated therein a nucleic acid construct coding for a fusion protein, which comprises at least a portion of a coat protein of the phage and a potential V_H binding-fragment, preferably spanning from upstream of an immunoglobulin heavy chain CDR1 to a position downstream of CDR3, said nucleic acid construct comprising a coding sequence which is biased to conserve part of the amino acid

constitution of said parental V_H binding-fragment at least part of which is a randomly or preferably a partially randomly generated variant (e.g. biased to preferred amino acid constitutions) of a segment of DNA encoding a CDR3 of a soluble non-camelid type parental V_H binding-fragment, such that said plurality of recombinant phage collectively encode fusion proteins having a plurality of different potential V_H binding-fragments related to said parental V_H binding-fragment, said fusion proteins being expressed in the absence of an immunoglobulin light chain protein or portions thereof on an outer surface of recombinant phage of the library, and wherein said potential V_H binding-fragments are adapted to be or capable of being expressed as soluble proteins and are preferably characterized by a CDR3 having 16 to 33 amino acids.

Preferably the potential V_H binding-fragments include the entire FR1 through to FR4 regions, although it is to be understood that partial deletions, particularly within CDR2, are contemplated to be within the scope of the invention.

Preferably, CDR3s of a variety of different lengths from 16 to 33 amino acids are predominantly represented among said potential V_H binding-fragments. Preferably CDR3s of a variety of different lengths, from 17 to 28 amino acids, or, from 17 to 23 amino acids are predominantly represented in the library. Although the term "predominant" ordinarily implies a majority representation of the specified long CDR3 potential V_H binding-fragments, the invention also contemplates an even less substantial representation, especially within a reasonably large size library ($>10^7$). Preferably, the specified long CDR3 potential V_H binding-fragments do have a majority representation within the library and more preferably an even greater or exclusive

representation.

A predominant representation of potential V_H binding-fragments having relatively long CDR3s of varying lengths, is preferably accomplished by randomly or partially randomly inserting varying numbers of nucleotide triplets in any part of a randomized portion of the parental V_H framework. This is accomplished by synthesizing primers of the desired length and nucleotide composition followed by PCR amplification. Desired randomization can also be achieved by biasing nucleotide composition of the said primer. This representation of long CDR3 potential binders may also be accomplished by pooling several libraries of potential V_H binding-fragments having randomized or partially randomized CDR3s of different respective uniform lengths. These strategies are not mutually exclusive.

Optionally, CDR3s of the same length as that of the parental V_H binding-fragment are predominantly or exclusively represented in said potential V_H binding-fragments.

In a preferred embodiment of the invention, the parental V_H binding-fragment is built on a human framework and preferably is the parental V_H binding-fragment identified in SEQ. ID. 1 which has a CDR3 of 23 amino acids in length. This parental V_H binding-fragment has preferred solubility characteristics which may be imparted to the potential V_H binding-fragments generated therefrom by biasing the amino acid constitution of the potential V_H binding-fragments in favour of the amino acid constitution of the parental V_H binding-fragment, so that the population of potential V_H binding-fragments as a whole will have a greater representation of higher solubility

binders.

Thus, the invention is preferably directed to a phage-display library comprising a plurality of recombinant phage, each of said recombinant phage having incorporated therein a nucleic acid construct coding for a fusion protein which comprises at least a portion of a coat protein of the phage and a potential V_H binding-fragment, preferably spanning from a position upstream of an immunoglobulin heavy chain CDR1, preferably including an entire FR1, to a position downstream of CDR3, preferably including an entire FR4, said fusion protein being expressed in the absence of an immunoglobulin light chain protein or portions thereof on an outer surface of recombinant phage of the library, said nucleic acid construct comprising a coding sequence which is biased to conserve part of the amino acid constitution of said parental V_H binding-fragment, at least part of which is a randomly or partially randomly generated variant of a segment of DNA encoding a CDR3 of a soluble non-camelid or non-camelid type parental V_H binding-fragment, such that said plurality of recombinant phage collectively encode a plurality of different potential V_H binding-fragments related to said parental V_H binding-fragment (preferably built on, adopted to, or adaptable to a human framework), and wherein the resulting population of potential V_H binding-fragments are as a whole adapted to be or capable of being expressed as soluble proteins. The resulting binding-fragment population of potential V_H binding-fragments preferably is characterized by a CDR3 having 16 to 33 amino acids, preferably 17 to 28 amino acids, alternatively 17-23 amino acids, more preferably 23 amino acids.

In another aspect, the invention is directed to a method of specifically retaining in the CDR3 region and other regions human sequence elements that confer favourable characteristics of intermolecular interaction and/or high solubility and to a phage display library having favourable characteristics of intermolecular interaction and/or high solubility, preferably when compared with potential V_H binding-fragments that have fully randomized hypervariable regions (particularly CDR3). In particular, we have found that favourable solubility and/or intermolecular interaction characteristics of a parental V_H binding-fragment can be maintained in the population of potential V_H binding-fragments in the course of randomizing the hypervariable regions by biasing all or selected amino acids residues to wild-type and/or biasing in favourfavour of amino acids residues that favour certain or a variety of types of intermolecular interaction. This is respectively accomplished by increasing the percentage amounts of nucleotide bases that represent wild-type amino acids and/or amino acids that provide favourable intermolecular interactions during the randomization procedure e.g. site directed PCR mutagenesis.

The selection of amino acids for randomisation or partial randomisation is based on adopting one or more of a variety of approaches including:

1. universal recognition of wild-type amino acids through a broad-based biasing of the wild-type amino acids in one or more regions of interest (approximately 10%-90% biasing) in order to maintain the characteristics of the parental V_H binding-fragment;
2. selective recognition of amino acids that are important to maintain as wild-type through biasing (approximately 10-100%) in order maintain conserved or strategic regions amino acid residues (e.g. G) of the parental V_H binding-fragment;

3. recognition of selected amino acids as important for intermolecular interaction and biasing those amino acids to wild-type, amino acids of the same type and/or other amino acids that are preferred for intermolecular interactions.

The invention encompasses a phage display library which is constructed using a parental V_H binding-fragment derived from the human parental type V_H binding-fragment identified in SEQ. ID. 1 or is built on any framework which is at least 80% homologous to the framework and other conserved regions of said fully human V_H chain. The invention also contemplates that said parental V_H binding-fragment, though not human, is adapted (e.g. humanised) or adaptable (to be adapted after selection of preferred binders) to a human framework.

In another embodiment, the invention also contemplates the random, biased or fixed occurrence of features disclosed in the camelid literature, for example pairable cysteines in CDR1 and CDR3 and/or the substitution of hydrophilic amino acids at least one of positions 44, 45, and 47 (Kabat numbering system).

In another separate aspect of the invention the parental V_H binding-fragment is sought to be reduced in size and the parental V_H binding-fragment is preferably modified by deleting a portion of the CDR2.

Brief Description of the Drawings

The invention will now be described with reference to the drawings, wherein:

- Figure 1 is a sequence diagram showing a preferred parental V_H binding-fragment according to the invention.
- Figure 2 is a sensogram overlay showing the binding characteristics of a potential V_H binding-fragment generated against anti-FLAG antibody (M2) using a phage display library according to the invention.
- Figure 3 is a diagrammatic representation of SJFI, the vector used to create the vector into which the library is cloned.
- Figure 4 is a listing of the nucleotide and amino acid sequence of A6 VH after introduction of the *Nhe*1 site.

Detailed Description of Preferred Embodiments

As shown in Figure 1, in a preferred embodiment the parental V_H binding-fragment is derived from a human IgM heavy chain, which comprises all or part of the FR1 and FR4 and spans the entire sequence therebetween. A partial sequence of this antibody BT32/A6 (A6) is disclosed in co-pending Canadian Application No. 2,192,079 (U.S. Patent No. 5,639,863), incorporated herein by reference. Missing sequences in the framework regions are supplied in Figure 1.

Accordingly, this human heavy domain binding-fragment is well suited for the development of a phage display library that is used to generate soluble binding-fragments that are useful for human diagnosis and therapy (limited HAMA response). The phage display libraries may be used to selectively generate molecular probesthat specifically interact with natural and synthetic molecules and macromolecules and can be used *in vitro* (i.e., a diagnostic) and *in vivo* (i.e., a diagnostic and/or therapeutic) as indicators, inhibitors and immunological agents. The types of natural and synthetic molecules and macromolecules include but are not limited to: antibodies

and fragments thereof; enzymes; cell receptors; proteins, polypeptides, peptides; polynucleotides, oligonucleotides; carbohydrates such as polysaccharides, oligosaccharides, saccharides; lipids; organic-based and inorganic-based molecules such as antibiotics, steroids, hormones, pesticides, herbicides, dyes, polymers.

Still referring to Figure 1, the framework and CDR regions are demarcated. The amino acid residue numbers in Figure 1 and throughout the disclosure refer to the Kabat numbering system (Kabat et al. 1991, Sequences of Proteins of Immunological Interest, publication No. 91-3242, U.S. Public Health Services, NIH, Bethesda MD) except in the sequence listings and where explicitly stated or otherwise implied. Concordance between the Kabat numbering system and the sequence listing is specified below:

With respect to a separate aspect of the invention in which the parental V_H binding-fragment is sought to be reduced in size, also demarcated is a region of CDR2 that may be deleted.

As shown in Figure 2, surface plasmon resonance analysis of an A6 library derived potential V_H binding fragment MR2-1 (-CYS) against anti-FLAG (M2) IgG demonstrates that this binding fragment had definitive on-rates and off-rates from which an equilibrium constant of 1.1 μM was calculated.

The following table charts particular amino acid residues or sequences of residues and preferred types of amino acid substitutions according to various embodiments of the invention with respect

to each of the various approaches to partial or complete randomisation described below:

1. universal recognition of wild-type amino acids through a broad based biasing, for example, most of the CDR3, (approximately 10%-90% on an individual amino acid basis) in favour of wild-type amino acids;

2. recognition of selected amino acids that are important to maintain as wild-type through biasing (approximately 10-100%) in order maintain conserved or strategic regions (e.g. residue) of the parental V_H binding-fragment;

3. recognition of selected amino acids as important for intermolecular interaction and biasing those amino acids to wild-type, amino acids of the same type or other amino acids that are preferred for intermolecular interactions.

Interpretation

Unless otherwise implied or stated, the term "biasing" and related forms of this term, are generally intended to refer to weighting in the course of introducing variation in the parental binding-fragment. Unless otherwise necessarily implied as a result of logistical considerations, it is to be understood that that the various embodiments which relate to choice of amino acids for random, biased or fixed substitution (specified in column 1) as well as the various embodiments related to types of substitutions (column 2) are not mutually exclusive. Moreover the various permutations and combinations of such substitutions are hereby disclosed. For example, substitutions referred to in row a. (any one or more amino acids and preferably all amino acids of residues 100a – 100h) #3 (at least approximately 50% wild-type amino acids) may combined with row b. (any one or more and preferably all of amino acids residues 100a, 100b, 100g and 100h) #2 (for instance, at least approximately 90% wild-type amino acids) so that, for instance,

any 3 of the amino acids in 100a – 100h are biased in favour of wild-type in approximately 50% of the potential V_H binding-fragments and 100a and 100b are biased in favour of wild-type in 90% of potential binding-fragments. By necessary implication the three amino acids that are biased in favour of wild-type are not residues 100a and 100b, but they may be any other three residues. Accordingly, the broadest possible interpretation is to be given to the disclosure of the various combinations and permutations of the embodiments disclosed herein. Furthermore, it is to be understood that each of the various embodiments described herein are disclosed, except insofar as logistically impossible, in reference to each of the various aspects and definitions of the invention. Moreover, it is to be understood that phrases such as at least approximately 10%, or approximately 10-100% are intended to specify a preference for each of the unit percentages between about 7 and 100% that are practically achievable by oligonucleotide primer design and PCR amplification described herein below, as well as other well known PCR techniques and techniques of Controlled mutation described in the art, and routine variations of such techniques. Furthermore, it is to be understood, for example, that 90% biasing in favour of wild-type amino acids at a given amino acid position is to be approximated by controlling the percentage amounts of each of the three relevant nucleotides (so that, for example, the product of the probabilities of occurrence of the three desired nucleotides in sequence in the growing chain is 90%) so as to supply 90% of correct coding triplet(s) and a total of 10% of random coding triplets, having regard to the degeneracy of the genetic code (for example if two different coding triplets result in a given amino acid, then the sum of the probabilities of achieving those two triplets will have to equal 90%). This is preferably accomplished on an amino acid by amino acid basis so that, for example the probability of achieving two and three wild-type amino acids in sequence, in the

case of 90% biasing is 0.81 and 0.73, respectively, etc. It is to be understood that this high level of biasing may be suitable only for part of the coding sequence into which variability is introduced and that higher levels of biasing are acceptable, when for example substantially all of the amino acids of a long CDR3 are biased, as disclosed in one of embodiments herein. Accordingly there is a balance to be struck between a large diverse library and biasing for multifactorial characteristics such as solubility. Nevertheless it is contemplated, in another aspect of invention that the final library may be a pooled library in which several libraries each having varying degrees of biasing to wild-type, for example, 60%, 50%, 40% and 30%, are pooled together to obtain the both desired variability and similarity. It is also to be understood that the preferred parental binding-fragment may be engineered to maximize the desired characteristic (e.g. solubility, intermolecular interaction) and then made the subject of libraries with varying degrees of biasing. In this connection, the library could be biased to be rich in amino acids, which are highly soluble. In a further aspect of the invention is to be understood that both arms (halves) of the preferred longer loop forming CDR3s may be biased to amino acids that are favoured for intermolecular interaction, preferably charged amino acids, so as to provide a method of generating, in addition to loop size, varying loop structures. This bias may be systematically introduced or systematically reduced by randomization, in cooperating pooled libraries having varying degrees of biasing.

It is to be understood that biasing of a percentage less than 100% implies unless otherwise implied or stated that the remaining percentage is fully randomized.

It is to be understood that conserved regions are those which are commonly found in at least other antibodies of the same type or in at least the same species of mammal.

All references named or referred to herein are all incorporated herein by reference.

It is to be understood that wild-type refers to the parental binding-fragment or to A6 wild-type parental binding-fragment depending on the context (A6 wild-type is not necessarily specified).

It is to be understood that the term "spanning" does not preclude deletions or additions within the parental V_H binding-fragment that are not inimical to the operation of the invention.

The term "camelid type" refers specifically to the camelid V_L interface.

The term "soluble" and "solubility" in reference to the preferred characteristics of a preferred parental binding-fragment of the invention and variants thereof, includes the generally ascribed meaning in the art and without limitation includes, (having regard to correlated phenomena) the relative amounts of naturally-folded recombinant protein released from the cell.

Unless otherwise specified, it is to be understood that a given "% biasing" or "% of binding-fragments" (or "biasing 10-100%", etc.) refers to biasing on an individual amino acid basis (though other techniques to accomplish the same effect might apparent to those skilled in the art).

Similarly, the specification that wild-type amino acids occur at a specified position or series of positions in, for example, at least approximately 50% of potential binding-fragments is intended to mean both that 50% biasing is sought at a given such position or that a total of 50% of the correct nucleotide triplets are represented. The use of the term "approximately" in reference to percentages is intended to accommodate attrition of various desired potential V_H binding-fragments, the inaccuracy of the assumption that the probabilistic outcomes will be achieved in practice and that certain variation in methods to accomplish the specified results is deemed to be suitable. The term 50% or approximately 50% in reference to an uneven number of amino acids residues means that either 1 more or 1 less than half of the amino acids is referred to. The term % homology or % homologous is deemed to include both of the following interpretations / methods of calculation: 1) an approximate percentage of the sequence referenced in terms of the number of common amino acids residues (e.g. 80% of 11 is understood to be an approximation insofar as application of the percentage does not yield a unit number of residues, in which case both the immediately higher number and immediately lower unit numbers, 9 and 8 respectively, are deemed to be covered); 2) the percentage of binding-fragments theoretically achievable that have the full wild-type seq., which is calculated as a product of the probabilities that the wild-type amino acid will occur at a given amino acid position.

Table 1

Amino Acid Residue #s	Description of Various Preferred Amino Acid Constitutions

a. At least one of 100a-100h, preferably at each position of 100a-100h	1. Randomize; 2. At least approximately 10% biasing in favour of wild-type amino acids; 3. At least approximately 50% biasing in favour of wild-type amino acids; 4. At least approximately 90% biasing in favour of wild-type amino acids; 5. Randomize, but bias 100f to wild-type (approximately 10-100%)
b. At least one amino acid of: 100a-100b and 100g-100h preferably at each position of 100a-100b and 100g-100h	1. Randomize; 2. Randomize with bias to wild-type (approximately 10-100%), preferably at least approximately 50% wild-type, alternatively at least approximately 90% wild-type amino acids; 3. Randomize with bias to one of the amino acids selected from the group consisting of tyrosine, histidine, glutamine, asparagine, lysine, aspartic acid and glutamic acid (approximately 10-100%)
c. At least one of 100b-100g, preferably at each position of 100b-100g	1. Randomize; 2. Delete;
d. 100a-100h	1. Random additions of up to 10 amino acids; 2. Random deletions of up to 7 amino acids;
e. 95-100o	1. Randomize; 2. Random additions of up to 10 amino acids; 3. Random deletions of up to 7 amino acids;
f. At least one of 95-100, preferably at each position of 95-100	1. Randomize; 2. Randomize with bias to wild-type (approximately 10-100%), preferably at least approximately 50% wild-type, or preferably at least approximately 90% wild-type amino acids; 3. Invariant (primer spans this region)
g. 101-102 conserved amino acids	1. Invariant (primer spans this region) 2. N/A
h. 100I-100o	1. Randomize 2. Randomize with bias to wild-type (approximately 10-100%), preferably at least approximately 50% wild-type, more preferably at least approximately 90% wild-type amino acids; 3. Randomize with bias to one of the amino acids selected from the group consisting of tyrosine, histidine, glutamine, asparagine,

	lysine, aspartic acid and glutamic acid (approximately 10-100%); 4. Randomize with bias to maintaining 100o as wild-type (10-100%).
i. At least one amino acid of 100a-100b, 100g-100h and 100l-100o, preferably at each position of 100a-100b, 100g-100h and 100l-100o	1. Randomize with bias to wild-type (approximately 10-100%), preferably at least approximately 50% wild-type, more preferably at least approximately 90% wild-type amino acids; 2. Randomize with bias to one of the amino acids selected from the group consisting of tyrosine, histidine, glutamine, asparagine, lysine, aspartic acid and glutamic acid (approximately 10-100%); 3. Bias to aromatic amino acids (10-100%)
j. 95-100h	1. Randomize but maintain any 5-10 consecutive amino acids as wild-type
k. 100a-100o	1. Randomize but maintain any 5-10 consecutive amino acids as wild-type

The practice of the present invention employs, unless otherwise indicated, conventional techniques of molecular biology (including recombinant techniques), microbiology, cell biology, biochemistry and immunology, which are within the skill of the art. Such techniques are explained fully in the literature, such as, "Molecular Cloning: A Laboratory Manual", second edition (Sambrook et al., 1989); "Oligonucleotide Synthesis" (M.J. Gait, ed., 1984); "Animal Cell Culture (R.I. Freshney, ed., 1987); "Methods in Enzymology" (Academic Press, Inc.); "Handbook of Experimental Immunology" (D.M. Wei & C.C. Blackwell, eds.); "Gene Transfer Vectors for Mammalian Cells" (J.M. Miller & M.P. Calos, eds., 1987); "Current Protocols in Molecular Biology (F.M. Ausubel et al., eds., 1987); "PCR: The Polymerase Chain Reaction", (Mullis et al., eds., 1994); "Current Protocols in Immunology" (J.E. Coligan et al., eds., 1991). These references are incorporated herein by reference. These techniques are applicable to the

production of the polynucleotides and polypeptides of the invention, and, as such, may be considered in making and practicing the invention. Particularly useful techniques for particular embodiments will be discussed in the sections that follow.

Recombinant genetic techniques have allowed cloning and expression of antibodies, functional fragments thereof and the antigens recognized. These engineered antibodies provide novel methods of production and treatment modalities. For instance, functional immunoglobulin fragments have been expressed in bacteria and transgenic tobacco seeds and plants. Skerra (1993) *Curr.Opin. Immunol.* 5:256:262; Fiedler and Conrad (1995) *Bio/Technology* 13:1090-1093; Zhang et al. (1993) *Cancer Res.* 55:3384-3591; Ma et al. (1995) *Science* 268:916; and, for a review of synthetic antibodies, see Barbas (1995) *Nature Med.* 1:836-839. These and more current references describing these techniques, which these references, particularly those well known to persons practicing in the relevant arts, are hereby incorporated herein by reference.

Nucleotide sequences can be isolated, amplified, and processed by standard recombinant techniques. Standard technique in the art include digestion with restriction nucleases, and amplification by polymerase chain reaction (PCR), or a suitable combination thereof. PCR technology is described in U.S. Patent Nos. 4,683,195; 4,800,159; 4,754,065; and 4,683,202, as well as *PCR: The Polymerase Chain Reaction*, Mullis et al., eds., Birkauswer Press, Boston (1994).

In addition to the specific PCR methods of biasing to wild-type A6 amino acid residues detailed

below, it is possible to produce multiple different oligonucleotide primers consisting of specified amino acid residues (one or more) of the wild-type A6 molecule (e.g. CDR3 residues), mixing these in appropriate concentrations with a completely randomized (e.g. CDR3) oligonucleotide primer and subjecting the mixture of oligonucleotide primers to PCR. This will result in a biased phage library population of one's choosing (i.e. the amounts of the selectively randomized and totally randomized primers in the mixture will determine the per cent of each CDR3 representation in the library).

Polynucleotides comprising a desired sequence can be inserted into a suitable vector, and the vector in turn can be introduced into a suitable host cell for replication and amplification. Polynucleotides can be introduced into host cells by any means known in the art. Cells are transformed by introducing an exogenous polynucleotide by direct uptake, endocytosis, transfection, f-mating or electroporation. Once introduced, the exogenous polynucleotide can be maintained within the cell as a non-integrated vector (such as a plasmid) or integrated into the host cell by standard methods. See, e.g., Sambrook et al. (1989). RNA can also be obtained from transformed host cell, or it can be obtained directly from the DNA by using a DNA-dependent RNA polymerase.

Suitable cloning and expression vectors include any known in the art, e.g., those for use in bacterial, mammalian, yeast and insect expression systems. Specific vectors and suitable host cells are known in the art and are not described in detail herein. See e.g. Gacesa and Ramji, *Vectors*, John Wiley & Sons (1994).

Phage display techniques are generally described or referenced in some of the preceding general references, as well as in U.S. Patent Nos. 4,593,002; 5,403,484; 5,837,500; 5,571,698; 5,750,373; 5,821,047; 5,223,409 and 5,702,892. "Phage Display of Peptides and Proteins", (Kay, Brian K. et al., 1996); "Methods in Enzymology", Vol. 267 (Abelson, John N., 1996); "Immunology Methods Manual", (Lefkovits, Ivan, 1997); "Antibody phage display technology and its applications", (Hoogenboom, Hennie R. et al., 1998). Immunotechnology 4 p.1-20

Generally, DNA encoding millions of variants of a parental binding-fragment can be batch-cloned into the phage genome as a fusion to the gene encoding one of the phage coat proteins (pIII, pVI or pVIII). Upon expression, the coat protein fusion will be incorporated into new phage particles that are assembled in the bacterium. Expression of the fusion product and its subsequent incorporation into the mature phage coat results in the ligand being presented on the phage surface, while its genetic material resides within the phage particle. This connection between ligand genotype and phenotype allows the enrichment of specific phage, e.g. using selection on immobilized target. Phage that display a relevant ligand will be retained, while non-adherent phage will be washed away. Bound phage can be recovered from the surface, reinfected into bacteria and re-grown for further enrichment, and eventually for analysis of binding. The success of ligand phage display hinges on the combination of this display and enrichment method, with the synthesis of large combinatorial repertoires on phage.

Creation of single chain Fv's is generally described in U.S. Patent Nos. 4,[illegible],778; 5,260,203; 5,482,858; 5,258,498 and 5,525,491.

EXAMPLE 1

Construction of single-domain A6-**based** (A6-**based** dAb) DNA templates: to facilitate construction of the A6-**based** dAb libraries, a *Nhe*I site (nucleotides underlined and bolded in Fig.1) was introduced at the amino acid residues 24-25 while maintaining the wild-type amino acid sequence. Briefly, the A6 VH gene was used as a PCR template to amplify a shorter internal fragment employing the primers A6VH/NheI - 5'(TGTTCAGCTAGCGGATTC)3' and A6VH/BstEII- 5'(TGAGGAGACGGTGACCGTTGTCCCTTGGCCCCAGATATCAAA)3'. These primers incorporate *Nhe*I and *BstE*II sites (underlined) at the 5' and 3' ends of the amplified product. PCR (polymerase chain reaction) was performed in a total volume of 50 µl containing 200 mM each of the four dNTPs, 100 pmol each of the two primers, 5 µl of 10X buffer (New England Biolabs (NEB)), and 2 units of Vent DNA polymerase (NEB). The amplified product was purified using QIAquick PCR Purification kit™ (QIAGEN, Mississauga, ON), digested with *Nhe*I and *BstE*II restriction endonucleases and subsequently ligated to the *Nhe*I/*BstE*II-restricted pSJF1-10A12 vector derived from pUC 8 (Narang et al., 1987) to replace a portion of the existing A6 VH gene; to construct the pSJF1 vector, the pUC 8 plasmid (Vierra and Messing, 1982; Messing, 1983) was modified by inserting the OmpA signal sequence and the His_5-carboxy tail between the EcoRI and HindIII restriction sites of the pUC 8 polylinker region, using oligonucleotide primers and PCR (Narang et al., 1987). Electro-competent *E.coli*

TG1 cells were prepared (Tung and Chow, 1995) and an aliquot of the ligated product was used to transform the cells. Transformation was carried out using the BIO-RAD Gene Pulser™ according to the manufacturer's instructions and the clone harbouring the **mutated** A6 dAb gene was confirmed by sequencing (Sanger, F. et al., 1977) using the AmpliTaq DNA Polymerase FS kit and 373A DNA Sequencer Stretch (PE Applied Biosystems, Mississauga, ON). All the cloning steps below were performed as previously described (Sambrook et al. 1989). The resulting vector is termed pSJF1-A6VH.*Nhe*I.

EXAMPLE 2

A6 dAb library construction: The steps of A6 dAb library construction involved a series of sequential PCR experiments.

(1) *Introduction of restriction sites to facilitate cloning*: To amplify the target DNA, the PCR mixture was first incubated at 95°C for 5 min, then subjected to 30 cycles of: 30 sec at 94°C, **1 min** at 40°C and, **1 min** at 72°C. The A6VH.*Nhe*I-containing plasmid, pSJF1-A6VH.*Nhe*I, was used as the template in PCR to amplify a shorter fragment using the primers A6VH/*Apa*II – 5'. (CATGACCACAGTGCACAGGAGGTCCAGC-TGCAGGAGTC) 3' and A6VH.FR3.F - 5' (TTTCACACAGTAATACAC) 3'. The PCR mixture contained 200 µM each of the four dNTPs, 0.2 pmol/µl each of the two primers, 1X buffer (Perkin Elmer), and 0.05 units/µl of AmpliTaq DNA polymerase (Perkin Elmer). (The former primer also introduces *Apa*II site at the 5' end of the PCR product.)

(2) *Randomization of the A6 dAb CDR3 residues*: The amplified fragment from step (1) was purified by QIAquick Gel Extraction kit™ (QIAGEN) and subsequently used as the template in a second PCR reaction using the primers A6VH.*Apa*II and A6VH.RndmCDR3.F - 5' (GCCCCAGATATCAAA20 [((A/C)NN)]TTTCACACAGTAATA)3'. At the protein level the second primer results in the randomization of the first 20 residues in CDR3. The PCR mixture was identical to above except that the concentration of the primers was increased to 0.5 pmol/µl to ensure that sufficient amounts of oligonucleotide primers and dNTPs were provided for the generation of a large randomized library.

(3) *Addition of a NotI restriction site, ligation to the phage vector and library construction*: The amplified fragments were purified as above and used as templates in a third round of PCR employing 2 pmol/ul each of the two primers A6VH/*Apa*II (described above) and A6VH.NotI.EXT.F – 5'(CGATTCTGCGGCCGCTGAGGAGACGGTGACCGTT-

GTCCCTTGGCCCCAGATA[redacted]AAA) 3'. (The latter primer incorporates the *Not*I site (underlined) at the 3' end of the amplified products.) The amplified fragments were purified using QIAquick PCR Purification kit™ (QIAGEN), digested with *Apa*I and *Not*I, and ligated to *Apa*I/*Not*I-digested fd-tet phage vector (McAfferty et al., 1990; Zacher et al., 1980). The ligated product was desalted using QIAquick PCR Purification kit™ (QIAGEN).

To determine the size of the library, immediately following the transformation and after the addition of the SOC medium (per L: bacto-tryptone, 20 g; bacto-yeast extract, 5 g; NaCl, 0.5 g; glucose, 3.6 g) a small aliquot of the electroporated cells were serially diluted in exponentially growing E. *coli* strain TG1 cells. Two hundred µl of the diluted cells were mixed with 3ml of 50°C top agar and immediately poured onto 2xYT (per L: bacto-tryptone, 16 g; bacto-yeast extract, 10 g; NaCl, 5 g) plates pre-warmed to 37°C. Plates were incubated overnight at 37°C and the number of plaques were used to determine the size of the library. Following this, the DNA inserts from single **plaques** were amplified using PCR. The size of the amplified product, determined by agarose gel electrophoresis, was used to determine the fraction of the library with full-sized A6 dAb inserts. Diversity of the library was determined to be in the range of 10^7-10^9.

The recombinant **phage** vectors, 1.5 µg, were mixed with 40 µl of competent *E. coli* strain TG1 and the cells were transformed by electroporation. Following transformation, 1 ml of SOC medium was added to each electroporation mixture (45 ml in total). The mixture was divided into three equal aliquots, each of which were added to tubes containing 3 ml of top agar at 50°C, vortexed immediately, poured onto pre-warmed 2xYT agar plates, and incubated at 37°C overnight. Five ml of sterile PBS (per L: NaCl, 8 g; KCl, 0.2 g; Na_2HPO_4, 1.44 g; KH_2PO_4, 0.24 g; pH 7.4) was added to the plates and the phage particles were eluted by gently shaking the plates at 4°C for 3 hr. The phage-containing PBS supernatant was collected, the plates rinsed

with an additional 5 ml of PBS and the two supernatants were pooled. The supernatants were centrifuged at 6000g for 15 min at 4°C, the cleared supernatant decanted and the phage were purified as described by Harrison et al. (1996). The phage pellet was dissolved in 20 ml of sterile PBS, divided into 100 µl aliquots and stored in liquid nitrogen.

EXAMPLE 3

Testing of the phage display A6 dAb library against the anti-FLAG M2 monoclonal IgG antibody:

The phage display A6-based dAb library was panned against the anti-FLAG M2 monoclonal antibody as described by the New England Biolabs (Beverly, MA) (NEB Technical Bulletin (1998): Ph.D.Ô Phage Display Peptide Library Kits; Knappik and Pluckthun, 1994); the FLAG peptide epitope recognized by the M2 monoclonal antibody is (X)YKXXD where the first position has a preference for aspartic acid (Miceli et al., 1994). On a random basis, considering the length of the randomized region of A6 CDR3 (i.e., 20 residues), the consensus sequence should occur at a frequency of $4x10^{-4}$. Thus, in the A6 dAb library with $2x10^{7}$ individual clones, the FLAG peptide epitope should be represented by approximately $5x10^{2}$ independent clones.

After three rounds of panning against M2 IgG thirty one clones from rounds two and three were selected and their A6 dAb genes sequenced. Twelve different A6 dAb genes with the FLAG consensus sequence were identified (Table 1).

EXAMPLE 4

Cloning and expression of A6 dAb genes

The A6 dAb genes were subcloned into an E. *coli* expression vector and the expressed A6 dAb proteins were isolated and purified to homogeneity. Briefly, 30 ml of LB containing 100 µg/ml ampicillin was inoculated with a single colony harboring the pSJF1-dAb and the culture was shaken at 240 rpm at 37°C overnight. In the morning the entire overnight culture was used to inoculate 1 liter of M9 medium supplemented with 5 µg/ml vitamin B1, 0.4% casamino acid and 100 ug/ml ampicillin. The culture was shaken at room temperature for 30 hr at 160 rpm and subsequently supplemented with 100 ml of 10x induction medium and 100 ul of 1M isopropylthiogalactoside (IPTG). The culture was shaken for another 60 hr, the periplasmic fraction was extracted by osmotic shock method (Anand et al., 1991), and the presence of A6 dAb in the extract was detected by Western blotting (MacKenzie 1994). The periplasmic fraction was dialyzed extensively in 10 mM HEPES (N-[2-hydroxyethyl]piperazine-N'-[2-ethanesulfonic acid]) buffer pH 7.0., 500 mM NaCl. The presence of the His_5 tag at the C-terminal of the A6 dAbs allowed a one step protein purification by immobilized metal affinity chromatography (IMAC) using HiTrap Chelating™ column (Phamacia). The 5-ml column was charged with Ni^{2+} by applying 30 ml of a 5 mg/ml $NiCl_2.6H_2O$ solution and subsequently washed with 15 ml deionized water. Purification was carried out as described (MacKenzie, 1994) except that the starting buffer was 10 mM HEPES buffer, 10 mM imidazole, 500 mM NaCl, pH 7.0, and the bound protein was eluted with a 10-500 mM imidazole gradient. The purity of the protein was determined by SDS-polyacrylamide gel electrophoresis. To detect the presence of dimer/multimer dAb in the protein preparation, gel filtration chromatography was performed using Superdex75 (Pharmacia) as described (Deng et al., 1995).

EXAMPLE 5

Affinity measurements: Kinetic binding parameters of the recombinant A6-derived dAb proteins derived from phage library panning of M2 IgG (Table 1) were characterized by surface plasmon resonance using a BIAcore biosensor system (Biocore AB Inc.) as described by Jonsson et al. (1991), Deng et al. (1995) and MacKenzie et al. (1996) . The M2 monoclonal IgG antibody was immobilized on the sensor chip and recombinant A6 dAb proteins were passed over the M2 IgG surface. In addition, the cysteine residue in the CDR3 of clone M2R2-1 was replaced with the parental A6 residue, Ser-100e; the generated clone, termed M2R2-1 (-Cys) (Table 1), was expressed, purified and analyzed by surface plasmon resonance as shown in Fig. 1. For the determination of rate constants, the sensorgram data presented in Fig.1 was fitted globally to a simple 1:1 interaction model using the BIAevaluation 3 software. The on-rate and off-rate were calculated to be 340 M^{-1} s^{-1} and 3.6 x 10^{-4} s^{-1}, respectively. Based on these kinetic constants, the dissociation equilibrium constant, K_d ,of the interaction is 1.1 μM.

Table 2. CDR3 sequences of A6-based dAb clones isolated by panning against M2 IgG. The FLAG consensus sequence is shown in bold.

A6 dAb clone	CDR3 sequence
M2R2-1	VQYGKHRRGSCIEVHP**EYKDFDI**
M2R2-2	NPPKPGAQARCVTTVK**DYKEFDI**
M2R2-4	AAIQTETARWCDRHPV**SYKMFDI**
M2R2-5	QTETQPLYNDCILRQA**GYKWFDI**

M2R2-9	MHTLQHYRNLCSYQLA**DYKHFDI**
M2R2-10	GLSGSRPNEQC**DYKTGDH**VQFDI
M2R2-13	LSGQNYTKTRCLVMQN**DYKMFDI**
M2R2-14	TAEPALSPQACMTKER**QYKDFDI**
M2R2-15	ETYMYTRGKYCRALSA**DYKLFDI**
M2R2-18	ESKASRTADQCSGPTP**GYKNFDI**
M2R3-4	GSQAIKNLSECLVRSD**DYKKFDI**
M2R3-13	GRYFQSKITSCENNDR**DYKLFDI**
M2R2-1 (-Cys)	VQYGKHRRGSSIEVHP**EYKDFDI**

EXAMPLE 6

Introducing genetic variation into the sequence corresponding to the A6 heavy chain CDR3 region: Oligonucleotides comprising randomly mutated CDR3 regions were prepared on an Applied Biosystems 394 DNA synthesizer as described above.

1. For 23 randomized residues (CDR3 1-23):

This anti-codon formula [(A/C)NN], is used and results in a reduction in possible codon usage from 64 to 32 and reduces the number of possible stop codons. Position one, therefore, comprises only A and C in the synthetic reaction mixture. For complete randomization of the second and third positions of the codons the dNTP mixture will comprise 25% each of A,G,C and T.

The 3' oligonucleotide randomizing primer was designed such that the last 15 nucleotides of framework 3 and the first 16 nucleotides of framework 4 were kept constant for hybridization. The nucleotides encoding the intervening amino acids, namely amino acids 1-23 of the CDR3 region were randomized using the following primer:
5' (GTTGTCCCTTGGCCCCA n[(A/C)NN]TTTCACACAGTAATA] 3' (Where n=23,. antisense strand).

Using a 50% A and 50% C for the first nucleotide position for each anti-codon triplet and 25% each of A, C, G, and T for the second and third nucleotide positions for n=23, complete randomization of the 23 amino acids of the A6 CDR3 is achieved.

2. For synthesis of CDRs comprising 15-23 residues:

The primers would be adapted by reducing n to 15-23 in the above primer formulae whilst keeping the flanking nucleotides constant.

3. For synthesis of CDR3s comprising 24-33 residues:

The primers would be adapted by increasing n to 24-33 in the above primer formula whilst keeping the flanking nucleotides constant.

4. Selective randomization:

For maintaining various percentages of wild type amino acid residues. This is achieved by creating residue substitutions by using different spiking levels of the various dNTPs as described below.

Example 6:

To achieve approximately 50% homology to wild type at any one position in the A6 dAb CDR3 region during antisense synthesis using the DNA synthesizer , the following example would be used. In the case of tyrosine, which is encoded by TAC or TAT (antisense strand GTA or ATA) the nucleotides would be spiked as follows for the antisense strand.

The nucleotide spiking levels would be as follows:

First anticodon nucleotide position: Only 80% of A and 20% of C is added, G and T are not added to reduce codon degeneracy.

For the second nucleotide position the mixture will consist of 80% T and approximately 6.67% of C, 6.67 of A and 6.67% of G.

For the third position the mixture will consist of 80% of A and approximately 6.67% of T and

6.67% of G and 6.7% C. The calculated probability of tyrosine would be 0.8x 0.8 x 0.8 x 100% = 51.2 %. Thus approximately 51% of the chains of the library will contain a wild-type A6 tyrosine in that specified position.

Example 7:

Using the same strategy in order to achieve approximately 50% homology to wild type at one or more positions the following example is useful.

Using only A and/or C in the first anticodon position the amino acid serine could have two codons these are AGT, TCT and TCG (antisense ACT, AGA and CGA, respectively).

The nucleotide spiking levels would be as follows:

First anticodon nucleotide position: 50% A and 50% C.

Second antisense nucleotide position: 35.35% C, 35.35% G, 14.65% A and 14.65% T

Third antisense nucleotide position: 35.35% A, 35.35% T, 14.65% C and 14.65% G.

The calculated probability of serine using this strategy is (1 x [0.3535+0.3535] x [0.3535+0.3535] x 100% = 50 %. Thus, approximately 50% of the chains will have a serine in that position.

Example 8:

To achieve approximately 10% homology to wild type at any one position in the A6 dAb CDR3 region during antisense synthesis using the DNA synthesizer, the following example would be

used. In the case of tyrosine which is encoded by TAC or TAT (antisense and GTA or ATA) the nucleotides would be spiked as follows for the anti sense strand. The nucleotide spiking levels would be as follows:

First anticodon nucleotide position: Only 47% of A and 53% of C is added; G and T are not added to reduce codon degeneracy.

For the second nucleotide position the mixture will consist of 47% T and approximately 17.67 % of C, 17.67 of A and 17.67% of G.

For the third position the mixture will consist of 47% of A and approximately 17.67% of T and 17.67% of G and 17.67% C. The calculated probability of tyrosine would be 0.47x 0.47 x 0.47 x 100% = 10.4 %. Thus approximately 10% of the chains of the library will contain a wild-type A6 tyrosine in that specified position.

Example 9:

To achieve approximately 90% homology to wild-type amino acids at any positions in the A6 dAb CDR3 region during antisense synthesis using the DNA synthesizer , the following example would be used. In the case of tyrosine which is encoded by TAC or TAT (antisense strand GTA or ATA) the nucleotides would be spiked as follows:

First anticodon nucleotide position: Only 97% of A and 3% of C is added, G and T are not added to reduce codon degeneracy. For this reason, only A and C are used in the first anticodon position for all 20 naturally occurring amino acids.

For the second nucleotide position the mixture will consist of 97% T and approximately 1 % of

C, 1% of A and 1% of G.

For the third position the mixture will consist of 97% of A and approximately 1% of T and 1% of G and 1% C. The calculated probability of tyrosine would be 0.97x 0.97 x 0.97 x 100% = 91.3 %. Thus approximately 90% of the chains of the library will contain a wild-type A6 tyrosine in that specified position.

Using the approaches in the examples above, approximately 10 % to approximately 90 % of wild type amino acid representation at one or more specified amino acid residues in the A6 CDR3 can be calculated and applied to the DNA synthesizer.

The present invention may be embodied in other specific forms without departing from the spirit or essential characteristics thereof. Certain adaptations and modifications of the invention will be obvious to those skilled in the art. Therefore, the presently discussed embodiments are considered to be illustrative and not restrictive. It is understood that the claims may refer to aspects or embodiments of the invention that are only inferentially referred to in the disclosure.

Other References:

Anand, N.N., Dubuc, G., Phip[redacted], J., MacKenzie, C.R., Sadowska, J., You[redacted], N.M., Bundle, D.R. and Narang, S.A. (1991). Synthesis and expression in *Escherichia coli* of cistronic DNA encoding an antibody fragment for a Salmonella serogroup B O-antigen. Gene **100**, 39-44.

Deng, S.-J., MacKenzie, C. R., Hirama, T., Brousseau, R., Lowary, T. L., Young, N. M., Bundle, D. R., and Narang, S. A. 1995. Analysis by surface plasmon resonance of the influence of valence on the ligand binding affinity and kinetics of an anti-carbohydrate antibody. Proc. Natl. Acad. Sci. U. S. A. **92**, 4992-4996

Jönsson, U., Fägerstam, L., Ivarsson, B., Johnsson, B., Karlsson, R., Lundh, K., Löfås, S., Persson, B., Roos, H., Rönnberg, I., Sjölander, S., Stenberg, E., Ståhlberg, R., Urbaniczky, C., Östlin, H., and Malmqvist, M. (1991). Real-time biospecific interaction analysis using surface plasmon resonance and a sensor chip technology. BioTechniques **11**, 620-627.

Knappik, A. and Pluckthun, A. (1994). An improved affinity tag based on the FLAG peptide for the detection and purification of recombinant antibody fragments. Biotechniques **17**, 754-761.

MacKenzie, C.R., Sharma, V., Brummel, D., Bilous, D., Dubuc, G., Sadowska, J., Young, N. M., Bundle, D. R., and Narang, S. A. 1994. Effect of $C_?$-$C_?$ domain switching on Fab activity and yield in Escherichia coli: synthesis and expression of genes encoding two anti-carbohydrate Fabs. Bio/Technology **12**: 390-395.

MacKenzie, C.R., Hirama, T., Deng, S.-J., Bundle, D.R., Narang, S.A., and Young. N.M. (1996). Analysis by surface plasmon resonance of the influence of valence on the ligand binding affinity and kinetics of an anti-carbohydrate antibody. J. Biol. Chem. **271**, 1527-1533

Miceli, R.M., Degraff, M.E. and Fischer, H.D. (1994). . J. Immunol. Methods **167**, 279-287.

Harrison, J.L, Williams, S.C., Winter, G. and Nissim, A. (1996). Screening of phage antibody libraries. Methods in Enzymology **267**, 83-109.

McAfferty, J., Griffiths, A.D., Winter, G. and Chiswell, D.J. (1990). .Nature **348**, 552-554.

Messing, J. (1983). New M13 vectors for cloning. Methods in Enzymology 101, 20-???.

Narang, S.A., Yao, F.-L., Michniewicz, J.J., Dubuc, G., Phipps, J. and Somorjai, R.L. (1987). Hierarchical strategy for protein folding and design: synthesis and expression of T4 lysozyme gene and two putative folding mutants. Protein Engineering **1**, 481-485.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual (2nd edition). Cold Spring Laboratory Laboratory, Cold Spring Harbor, NY.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA, **74**, 5463-????.

Tung, W.L. and Chow, K.-C. (1995). A modified medium for efficient transformation of *E.coli*. Trends in Genetics **11**, 128-129.

Vierra, O. and Messing, J. (1982). The pUC plasmids , an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19**, 259-268.

Zacher, A.N., Stock, C.A., Golden, J.W. and Smith, G.P. (1980). . Gene **9**, 127-140.

THE EMBODIMENTS OF THE INVENTION IN WHICH AN EXCLUSIVE PROPERTY OR PRIVILEGE IS CLAIMED ARE DEFINED AS FOLLOWS:

We claim:

1. A heterogeneous population of replicable genetic packages which collectively display a plurality of different V_H binding-fragments, each said genetic package being genetically alterable and having an outer surface including a genetically determined outer surface protein, each package including a nucleic acid construct coding for a fusion protein which comprises at least a portion of said outer surface protein and a potential V_H binding-fragment spanning from upstream of an immunoglobulin heavy chain CDR1 to a position downstream of CDR3, said nucleic acid construct comprising a CDR3 coding sequence at least part of which is a randomly or partially randomly generated variant of a segment of DNA encoding a CDR3 of a non-camelid or a non-camelid type parental V_H

THE EMBODIMENTS OF THE INVENTION IN WHICH AN EXCLUSIVE PROPERTY OR PRIVILEGE IS CLAIMED ARE DEFINED AS FOLLOWS:

We claim:

1. A heterogeneous population of replicable genetic packages which collectively display a plurality of different V_H binding-fragments, each said genetic package being genetically alterable and having an outer surface including a genetically determined outer surface protein, each package including a nucleic acid construct coding for a fusion protein which comprises at least a portion of said outer surface protein and a potential V_H binding-fragment spanning from upstream of an immunoglobulin heavy chain CDR1 to a position downstream of CDR3, said nucleic acid construct comprising a CDR3 coding sequence at least part of which is a randomly or partially randomly generated variant of a segment of DNA encoding a CDR3 of a non-camelid or a non-camelid type parental V_H

binding-fragment such that said heterogeneous population of replicable genetic packages collectively encode fusion proteins having a plurality of different potential V_H binding-fragments related to said parental V_H binding-fragment, said fusion proteins being expressed in the absence of an immunoglobulin light chain protein or portions thereof on an outer surface of said genetic packages, and wherein said potential V_H binding-fragments are adapted to be or capable of being expressed as soluble proteins and are characterized by a CDR3 having 16 to 33 amino acids.

1'. A heterogenous population of replicable genetic packages as claimed in claim 1, wherein said potential V_H binding-fragments comprise the entirety of FR1 to FR4.

1a. A heterogenous population of replicable genetic packages as claimed in claim 1, wherein each said replicable genetic packages is a phage and said heterogeneous population of replicable genetic packages collectively constitute a phage display library.

1b. A phage-display library, comprising:

a plurality of recombinant phage, each of said recombinant phage having incorporated therein a nucleic acid construct coding for a fusion protein which comprises at least a portion of a coat protein of the phage and a potential V_H binding-fragment spanning from upstream of an immunoglobulin heavy chain CDR1 to a position downstream of CDR3, said nucleic acid construct comprising a CDR3 coding sequence at least part of which is a randomly or partially randomly generated variant of a segment of DNA encoding a CDR3 of a non-camelid parental heavy chain binding-fragment such that said plurality of recombinant phage collectively encode fusion proteins having a plurality of different potential V_H binding-fragments related to said parental heavy chain binding-fragment, said fusion proteins being expressed in the absence of an immunoglobulin light chain protein or portions thereof on an outer surface of recombinant phage of the library, and wherein said potential V_H binding-fragments are adapted to be or capable of being expressed as soluble proteins and are characterized by a CDR3 having 16 to 33 amino acids.

1b' A phage display library as claimed in claim 1b, wherein said potential V_H binding-fragments comprise the entirety of FR1 to FR4.

1b'' A phage-display library comprising a plurality of recombinant phage, each of said recombinant phage having incorporated therein a nucleic acid construct coding for a fusion protein which comprises at least a portion of a coat protein of the phage and a potential V_H binding-fragment spanning from a position upstream of an immunoglobulin heavy chain CDR1, including FR1, to a position downstream of CDR3, including FR4, said nucleic acid construct comprising a CDR3 coding sequence at least part of which is a randomly or partially randomly generated variant of a segment of DNA encoding a CDR3 of a soluble human parental V_H binding-fragment and one or more segments of

coding sequence which corresponds to the sequence of said such that said plurality of recombinant phage collectively encode fusion proteins having a plurality of different human potential V_H binding-fragments related to said parental V_H binding-fragment, said fusion proteins being expressed in the absence of an immunoglobulin light chain protein or portions thereof on an outer surface of recombinant phage of the library, and wherein the resulting population of potential V_H binding-fragments are **as a whole** adapted to be or capable of being expressed as soluble proteins and are characterized by a CDR3 having 17 to 23 amino acids.

2. A phage display library as claimed in claim 1a or 1b, wherein CDR3s of a variety of different lengths from 16 to 33 amino acids are predominantly represented in said potential heavy chain binding-fragments.

3. A phage display library as claimed in claim 1a or 1b, wherein CDR3s of a variety of different lengths from 17 to 23 amino acids are predominantly represented in said potential heavy chain binding-fragments.

4. A phage display library as claimed in claim 1a, wherein CDR3s of 23 amino acids in length are predominantly represented in said potential heavy chain binding-fragments.

5. A phage display library as claimed in claim 1a or 1b, wherein said potential V_H binding-fragment is built on a human framework.

5. a) A phage display library as claimed in claim 1a or 1b, wherein said parental heavy chain binding-fragment is derived from a human V_H chain identified in SEQ. ID. 1 or is built on any framework which is at least 80%homologous to the framework and other conserved regions of said human V_H chain.

6. A phage display library as claimed in claim 1a or 1b, wherein said parental V_H binding-fragment is adapted or adaptable to a human framework.

6'. A phage display library as claimed in any of the preceding claims, wherein the amino acids in one or more series of CDR3 amino acids selected from the groups of amino acids consisting of 95-100 and 100i-100n are preserved in approximately at least 90% or approximately 100% of said potential V_H binding-fragments.

6''. A phage display library as claimed in any of the preceding claims, wherein one or more amino acids in one or more series of CDR3 amino acids selected from the groups of amino acids consisting of 95-100, 100i-100n, 100o-102 and 101-102 are preserved, on an

amino acid by amino acid basis, in approximately at least 90% or approximately 100% of said potential V_H binding-fragments.

7. A phage display library as claimed in any of claims 1a to 6, wherein said potential heavy chain binding-fragments have a native human V_L interface at positions 44, 45, and 47

7. a) A phage display library as claimed in any of claims 1a to 6, wherein said potential heavy chain binding-fragments have non-hydrophobic amino acids at least one of positions 44, 45, and 47.

8. A phage-display library as claimed in claim 7 or 7a, wherein said potential heavy chain binding-fragments are further characterized by a CDR3 having a partial amino acid sequence which is at least 90% homologous to at least one region of conserved amino acids selected from those regions identified in Figure 1.

9. A phage-display library as claimed in claim in any of claims 1 to 8, wherein binding-fragment at least approximately 50% of the amino acids in amino acid sequence 100a-100h shown in Figure 1 are biased in favour of wild-type A6 in at least 10% of said potential V_H binding-fragments.

10. A phage-display library as claimed in claim 8, wherein at least approximately 50% of said potential heavy chain binding-fragments at least approximately 90% of the amino acid residues in sequence 100a-100h shown in Figure 1 are biased to wild-type.

11. A phage-display library as claimed in claim 9 or 10, wherein or more individual amino acids in positions 100a-100b and 100g-100h, or 100a–100c and 100f-100h, are wild-type in at least approximately 10% of said potential heavy chain binding-fragments.

12. A phage-display library as claimed in claim 9 or 10, wherein individual amino acids in positions 100a-100b and 100g-100h, or 100a–100c and 100f-100h, are wild-type in at least approximately 50% of said potential heavy chain binding-fragments.

13. A phage-display library as claimed in claim 8, wherein at least 50% of individual amino acids in positions 95-100 are biased in favour of wild type in at least 10% of said potential binding-fragments.

14. A phage-display library as claimed in claim 8, wherein at least 90% of the individual amino acids in positions 95-100 of Figure 1 are biased in favour of wild-type at least 10% of said potential binding-fragments.

15. A phage-display library as claimed in claim 8, 13 or 14, wherein at least 50% of the individual amino acids in positions 100*i* - 100*n* in Figure 1 are biased in favour of wild-type in at least approximately 10% of said potential binding-fragments.

16. A phage-display library as claimed in any of claims 8, 13 or 14, wherein at least 50% of the individual amino acids in positions 100*i* - 100*n* in Figure 1 are biased in favour of wild-type in at least approximately 50% of said potential binding-fragments.

17. A phage-display library as claimed in any of claims 1 to 16, wherein individual amino acids in any one or more of positions 100a-100b, 100g-100h, 100l and 100o are biased to be wild-type, aromatic or selected exclusively from the group consisting of tyrosine, histidine, glutamine, asparagine, lysine, aspartic acid and glutamic acid, in at least approximately 10% of said potential binding-fragments.

18. A phage-display library as claimed in any of claims 1 to 16, wherein amino acids in any one or more of positions 100a-100b, 100g-100h, 100l and 100o are biased to be wild-type, aromatic or selected exclusively from the group consisting of tyrosine, histidine, glutamine, asparagine, lysine, aspartic acid and glutamic acid, in at least approximately 50% of said potential binding-fragments.

19. A phage-display library as claimed in any of preceding claims 1 to 8, wherein at least 5 consecutive amino acids among residues 95-100n shown in SEQ. ID. 1 are biased to be the wild-type amino acids in at least approximately 10% of said potential heavy chain binding-fragments.

20. A phage-display library as claimed in any of preceding claims 1 to 8, wherein at least 8 consecutive amino acids among residues 95-100n shown in SEQ. ID. 1 are biased to be the wild-type amino acids in at least approximately 10% of said potential heavy chain binding-fragments.

21. A phage-display library as claimed in any of preceding claims 1 to 8, wherein at least 10 consecutive amino acids among residues 95-100n shown in SEQ. ID. 1 are biased to be the wild-type amino acids in at least approximately 10% of said potential heavy chain binding-fragments.

22. A phage-display library as claimed in any of the preceding claims, wherein at least amino acids positions 100a-100b to 100f-100h and 100m are biased to be wild-type amino acids in at least approximately 50% of said potential binding-fragments.

22. a) A phage-display library as claimed in any of the preceding claims, wherein at least amino acids positions 100f to 100m are biased to be wild-type amino acids in at least approximately 50% of said potential binding-fragments.

23. A phage-display library as claimed in any of the preceding claims, wherein at least amino acids positions 100o to 102 or 101 to 102 are biased to be wild-type amino acids.

24. A phage-display library as claimed in any of the preceding claims, wherein framework regions are at least approximately 90% homologous to that of the wild-type parental binding-fragment shown in SEQ. ID. 1.

25. A phage-display library as claimed in any of the preceding claims, wherein the CDR2 region is at least approximately 80% homologous to that of the wild-type parental binding-fragment shown in SEQ. ID. 1.

26. A phage-display library as claimed in any of the preceding claims, wherein the CDR1 region is at least approximately 80% homologous to that of the wild-type parental binding-fragment shown in SEQ. ID. 1.

26'. A phage display library as claimed in any of the preceding claims, wherein the CDR1 region is biased to have a cysteine residue for possible loop forming interaction with any randomly generated cysteine residues in CDR3.

27. 27a. phage-display library as claimed in any of preceding claims 1b-26', wherein said recombinant phage are constructed in an M-13 derived vector and said phage coat protein is pIII.

28. A population of genetic packages having a genetically determined outer surface protein including genetic packages which collectively display a plurality of potential binding-fragments in association with said outer surface protein, each package including a nucleic acid construct coding for a fusion protein, which is at least a portion of said outer surface protein and a variant of at least one soluble non-camelid or non-camelid type of parental binding-fragment, wherein at least part of said construct encodes or is biased in favour of encoding the amino acid constitution of said parental binding-fragment such that said plurality of different potential heavy chain binding domains are on the whole adapted to be or better capable of being expressed as soluble proteins.

29. A population of genetic packages having a genetically determined outer surface protein including genetic packages which collectively display a plurality of potential binding-fragments in association with said outer surface protein, each package including a nucleic acid construct coding for a fusion protein which is at least a portion of said outer surface protein and a variant of at least one soluble type of parental binding-fragment, wherein at least part of said construct is only partly randomized in that it is biased in favour of

encoding the amino acid constitution of said parental binding-fragment such that said plurality of different potential heavy chain binding domains are on the whole adapted to be or better capable of being expressed as soluble proteins.

30. A population of genetic packages having a genetically determined outer surface protein including genetic packages which collectively display a plurality of potential V_H binding-fragments in association with said outer surface protein, each package including a nucleic acid construct coding for a fusion protein which is at least a portion of said outer surface protein and a variant of at least one soluble parental V_H binding-fragment, wherein at least part of said construct including at least part of CDR3 (other than residues 101-102) encodes or is biased in favour of encoding the amino acid constitution of said parental V_H binding-fragment such that said plurality of different potential heavy chain binding domains are on the whole adapted to be or better capable of being expressed as soluble proteins.

31. A population of genetic packages having a genetically determined outer surface protein including genetic packages which collectively display a plurality of potential binding-fragments in association with said outer surface protein, each package including a nucleic acid construct coding for a fusion protein which is at least a portion of said outer surface protein and a variant of at least one soluble parental immunoglobulin binding-fragment, wherein at least part of said construct including a part of CDR3 which is randomized to create variation among said potential binding-fragments is biased in favour of encoding the amino acid constitution of said parental V_H binding-fragment such that said plurality of different potential heavy chain binding domains are on the whole adapted to be or better capable of being expressed as soluble proteins.

32. A population of genetic packages a claimed in any of claims 28, 29 and 31, wherein said genetic package is a phage and said soluble parental binding-fragment is selected from the group consisting of an scFv, Fab, V_H, Fd, Fabc, $F(ab')_2$, $F(ab)_2$ derived from A6.

33. A population of genetic packages or phage as claimed in any of preceding claims comprising a plurality of libraries, which are pooled, wherein at least a first and second of said pooled libraries differ in the degree of biasing to wild-type amino acids.

34. A population of genetic packages or phage as claimed in any of preceding claims comprising a plurality libraries which are pooled, wherein at least a first and second of said pooled libraries differ in the percentage biasing in favour of amino acids, preferably within CDR3, which have relatively high solubility characteristics, e.g. polar or charged amino acids.

35. A population of genetic packages of phage as claimed in any of the preceding claims, comprising a plurality of libraries which are pooled, wherein at least a first and second of such pooled libraries differ in degree of biasing for amino acid that are preferred for

intermolecular interaction, including but not limited to, tyrosine, histidine, glutamine, asparagine, lysine, aspartic acids and glutamic acid.

FIGURE 1

Structure of V_H domain of human A6 antibody.

1	2	3	4	5	6	7	8	9	10	11	12	13	14
GAG	GTC	CAG	CTG	CAG	GAG	TCT	GGG	GGA	GGC	TTA	GTC	CAG	CCT
E	**V**	**Q**	**L**	**Q**	**E**	**S**	**G**	**G**	**G**	**L**	**V**	**Q**	**P**

15	16	17	18	19	20	21	22	23	24	25	26	27	28
GGG	GGG	TCC	CTG	AGA	CTC	TCC	TGT	TCA	GCT	AGC	GGA	TTC	ACC
G	**G**	**S**	**L**	**R**	**L**	**S**	**C**	**S**	**A**	**S**	**G**	**F**	**T**

29	30	31	32	33	34	35	36	37	38	39	40	41	42
TTC	AGT	AGC	TAT	GCT	ATG	CAC	TGG	GTC	CGC	CAG	GCT	CCA	GGG
F	**S**	**S**	**Y**	**A**	**M**	**H**	**W**	**V**	**R**	**Q**	**A**	**P**	**G**
		CDR1											

43	44	45	46	47	48	49	50	51	52	a	53	54	55
AAG	GGA	CTG	GAA	TAT	GTT	TCA	GCT	ATT	AGT	AGT	AAT	GGG	GGT
K	**G**	**L**	**E**	**Y**	**V**	**S**	**A**	**I**	**S**	**S**	**N**	**G**	**G**
							CDR2						

56	57	58	59	60	61	62	63	64	65	66	67	68	69
AGC	ACA	TAC	TAC	GCA	GAC	TCC	GTG	AAG	GGC	AGA	TTC	ACC	ATC
S	**T**	**Y**	**Y**	**A**	**D**	**S**	**V**	**K**	**G**	**R**	**F**	**T**	**I**

70	71	72	73	74	75	76	77	78	79	80	81	82	a
TCC	AGA	GAC	AAT	TCC	AAG	AAC	ACT	CTG	TAT	CTT	CAA	ATG	AGC
S	**R**	**D**	**N**	**S**	**K**	**N**	**T**	**L**	**Y**	**L**	**Q**	**M**	**S**

b	c	83	84	85	86	87	88	89	90	91	92	93	94
AGT	CTG	AGA	GCT	GAG	GAC	ACG	GCT	GTG	TAT	TAC	TGT	GTG	AAA
S	**L**	**R**	**A**	**E**	**D**	**T**	**A**	**V**	**Y**	**Y**	**C**	**V**	**K**

95	96	97	98	99	100	a	b	c	d	e	f	g	h
GAC	AGG	TTA	AAA	GTG	GAG	TAC	TAT	GAT	AGT	AGT	GGT	TAT	TAC
D	**R**	**L**	**K**	**V**	**E**	**Y**	**Y**	**D**	**S**	**S**	**G**	**Y**	**Y**

CDR3

i	j	k	l	m	n	o	101	102	103	104	105	105	107
GTT	TCT	CGG	TTC	GGT	GCT	TTT	GAT	ATC	TGG	GGC	CAA	GGG	ACA
V	**S**	**R**	**F**	**G**	**A**	**F**	**D**	**I**	**W**	**G**	**Q**	**G**	**T**

108	109	110	111	112	113
ACG	GTC	ACC	GTC	TCA	TCA
T	**V**	**T**	**V**	**S**	**S**

FIGURE 2

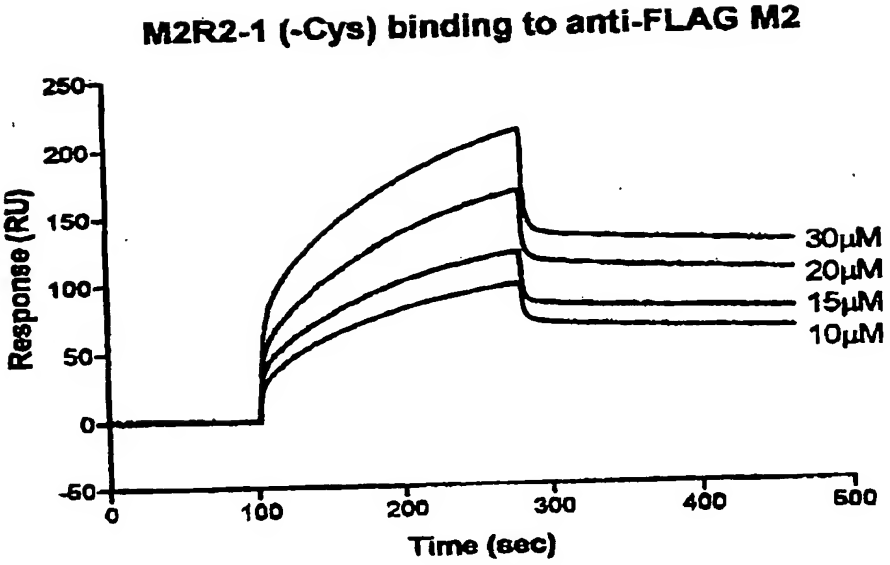
M2R2-1 (-Cys) binding to anti-FLAG M2
Response (RU)
250
200
150
100
50
0
-50
0
100
200
300
400
500
Time (sec)
30μM
20μM
15μM
10μM

Figure 3. pSJF1 expression vecto[redacted]1 was derived from the pUC8 cloning vector, which contains a Col[redacted] of replication 867, the LacZ operon with an intervening M13mp8 polylinker and an ampicillin resistance gene (see Appendix I for pUC8 components and seq[redacted] To construct pSJF1, the OmpA signal sequence and the His_6 tail were inserted between the EcoRI and Hind III restriction sites of the pUC8 polylinker region, using oligonucleotide primers and PCR. To change the antibiotic resistance from ampicillin to kanamycin, the ampicillin resistance gene was interrupted by digesting with *Sca*I, which results in a blunt cut within the ampicillin resistance gene. The kanamycin resistance gene from the pUC-4K vector (Pharmacia Biotech, Uppsala, Sweden) was inserted by first isolating the coding region with a *Hinc*II digestion of the pUC-4K vector and then ligating the blunt ends of the *Hinc*II digestion with the open vector.

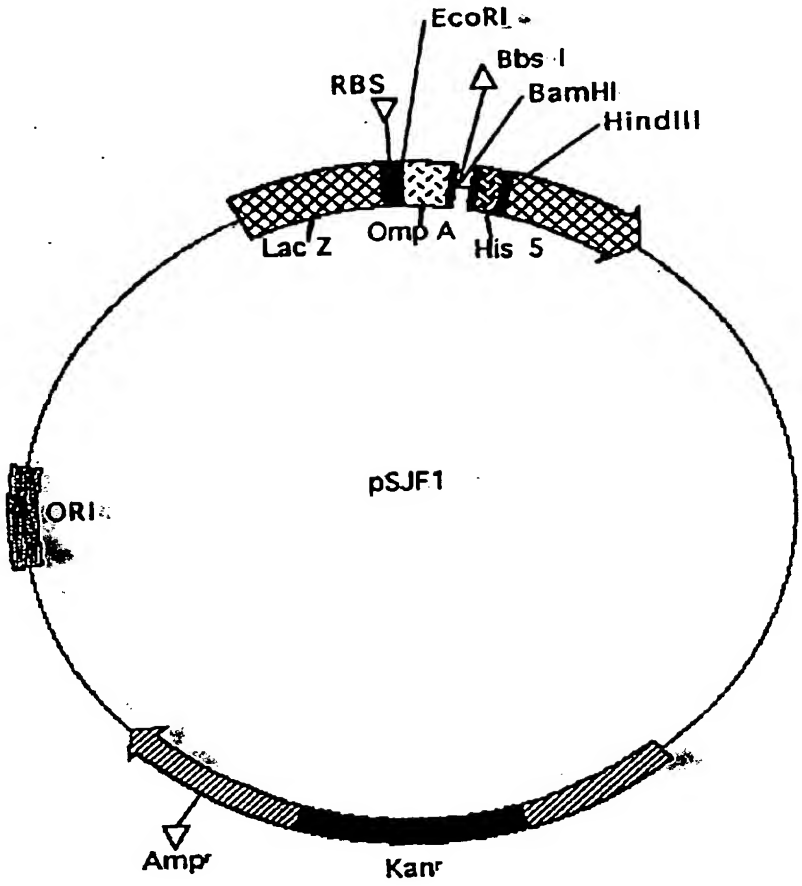

FIGURE 4

1 2 3 4 5 6 7 8 9 10 11 12 13 14

GAG GTC CAG CTG CAG GAG TCT GGG GGA GGC TTA GTC CAG CCT

E V Q L Q E S G G G L V Q P

15 16 17 18 19 20 21 22 23 24 25 26 27 28

GGG GGG TCC CTG AGA CTC TCC TGT TCA GCT AGC GGA TTC ACC

G G S L R L S C S A S G F T

29 30 31 32 33 34 35 36 37 38 39 40 41 42

TTC AGT AGC TAT GCT ATG CAC TGG GTC CGC CAG GCT CCA GGG

F S S Y A M H W V R Q A P G

CDR1

43 44 45 46 47 48 49 50 51 52 a 53 54 55

AAG GGA CTG GAA TAT GTT TCA GCT ATT AGT AGT AAT GGG GGT

K G L E Y V S A I S S N G G

CDR2

56 57 58 59 60 61 62 63 64 65 66 67 68 69

AGC ACA TAC TAC GCA GAC TCC GTG AAG GGC AGA TTC ACC ATC

S T Y Y A D S V K G R F T I

70 71 72 73 74 75 76 77 78 79 80 81 82 a

TCC AGA GAC AAT TCC AAG AAC ACT CTG TAT CTT CAA ATG AGC

S R D N S K N T L Y L Q M S

b c 83 84 85 86 87 88 89 90 91 92 93 94

AGT CTG AGA GCT GAG GAC ACG GCT GTG TAT TAC TGT GTG AAA

S L R A E D T A V Y Y C V K

95 96 97 98 99 100 a b c d e f g h

GAC AGG TTA AAA GTG GAG TAC TAT GAT AGT AGT GGT TAT TAC

D R L K V E Y Y D S S G Y Y

CDR3

i j k l m n o 101 102 103 104 105 105 107

GTT TCT CGG TTC GGT GCT TTT GAT ATC TGG GGC CAA GGG ACA

V S R F G A F D I W G Q G T

108 109 110 111 112 113

ACG GTC ACC GTC TCA TCA

T V T V S S

Figure 4 is a listing of the nucleotide and amino acid sequence of A6 VH after introduction of the *Nhe*1 site.